Ubinet, Distributed Optimization and Games 2015-2016	December 9, 2015
Lecture 2	
Lecturer: Giovanni Neglia	*Scribe: Giovanni Neglia*

NOTE: The content of these notes has not been formally reviewed by the lecturer. It is recommended that they are read critically.

In this lesson we provide some background on convex sets, convex functions and the Lagrange multipliers method for convex optimization problems.

1 Convexity

Definition 1 (Convex set)**.** *A subset $C \subset \mathbb{R}^n$ is convex if for each $\mathbf{x}, \mathbf{y} \in C$, and $\alpha \in [0, 1]$, it holds that*

$$\alpha\mathbf{x} + (1 - \alpha)\mathbf{y} \in C.$$

Definition 2 (Convex function)**.** *A function $f : C \to \mathbb{R}$, where C is a convex set, is convex if*

$$f(\alpha\mathbf{x} + (1 - \alpha)\mathbf{y}) \leq \alpha f(\mathbf{x}) + (1 - \alpha)f(\mathbf{y}),$$

for each $\mathbf{x}, \mathbf{y} \in C$, and $\alpha \in [0, 1]$.

Definition 3 (Concave function)**.** *A function $f : C \to \mathbb{R}$, where C is a convex set, is concave if $-f$ is convex, or equivalently if*

$$f(\alpha\mathbf{x} + (1 - \alpha)\mathbf{y}) \geq \alpha f(\mathbf{x}) + (1 - \alpha)f(\mathbf{y}),$$

for each $\mathbf{x}, \mathbf{y} \in C$, and $\alpha \in [0, 1]$.

2 Lagrange multipliers

Given a function $f : \mathbb{R}^n \to \mathbb{R}$ consider the optimization problem

$$\begin{aligned} \underset{\mathbf{x}\in\mathbb{R}^n}{\text{minimize}} \quad & f(\mathbf{x}) \\ \text{subject to} \quad & h_i(\mathbf{x}) = 0, \ i = 1, \dots, m \\ & g_j(\mathbf{x}) \leq 0, \ j = 1, \dots, r \end{aligned} \tag{1}$$

Definition 4 (Lagrangian)**.** *The* Lagrangian function $L : \mathbb{R}^{n+m+r} \to \mathbb{R}$ *is defined as*

$$L(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) = f(\mathbf{x}) + \sum_{i=1}^{m} \lambda_i h_i(\mathbf{x}) + \sum_{j=1}^{r} \mu_j g_j(\mathbf{x}) \tag{2}$$

where $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \dots \lambda_m)^T$ is the vector of Lagrange multipliers *corresponding to the equality constraints, and $\boldsymbol{\mu} = (\mu_1, \mu_2, \dots \mu_r)^T$ is the vector of Lagrange multipliers corresponding to the inequality constraints.*

The following theorem is the tool that allows us to solve most of the network optimization problems discussed in our course.

Theorem 1 (Convex optimization with linear constraints)**.** *Consider the optimization problem*

$$\begin{aligned} \underset{\mathbf{x}\in\mathbb{R}^n}{\text{minimize}} \quad & f(\mathbf{x}) \\ \text{subject to} \quad & \mathbf{a}_i^T\mathbf{x} = b_i, \ i = 1, \dots, m \\ & \mathbf{c}_j^T\mathbf{x} \leq d_j, \ j = 1, \dots, r \\ & \mathbf{x} \in X, \end{aligned} \tag{3}$$

where f is convex and continuously differentiable and X is a polyhedral set, i.e. a set specified by the intersection of a finite number of closed half spaces. Then $\mathbf{x}^*$ is a global minimum for this optimization problem if and only if there exist λ_i^*, $i \in \{1, 2, \dots m\}$ and μ_j^*, $j \in \{1, 2, \dots r\}$ such that:

1. *$\mathbf{x}^*$ is feasible,*
2. *$\mu_j^* \geq 0$ and $\mu_j^*(\mathbf{c}_j^T \mathbf{x}^* - d_j) = 0$, $\forall j \in \{1, 2, \dots r\}$,*
3. *$\nabla_{\mathbf{x}} L(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*)^T (\mathbf{x} - \mathbf{x}^*) \geq 0, \ \ \forall \mathbf{x} \in X$.*

Proof: The result can be easily derived adapting the proof of Theorem 3.4.1 in [1]. □

Remark 1 (Complementary slackness)**.** *The conditions $\mu_j^*(\mathbf{c}_j^T \mathbf{x}^* - d_j) = 0$ are called* complementary slackness conditions*, because every time the constraint $\mathbf{c}_j^T \mathbf{x}^* - d_j \leq 0$ is slack (meaning that $\mathbf{c}_j^T \mathbf{x}^* - d_j < 0$), the constraint $\mu_j^* \geq 0$ must not be slack (meaning that $\mu_j^* = 0$).*

Remark 2 (Flexible assignment of Lagrange multipliers)**.** *We observe that the polyhedral set X can be expressed by a set of linear inequalities analogous to those explicitly considered in problem* (3)*. Theorem 1 allows us then to include an arbitrary set of constraints in the Lagrangian and take into account the other constraints through the set X.*

Remark 3 ($X = \mathbb{R}^n$)**.** *If the set X coincides with $\mathbb{R}^n$, i.e. all the constraints are explicitly taken into account through the Lagrangian, then the condition $\nabla_{\mathbf{x}} L(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*)^T (\mathbf{x} - \mathbf{x}^*) \geq 0, \ \ \forall \mathbf{x} \in \mathcal{R}^n$ leads to*

$$\nabla_{\mathbf{x}} L(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*) = 0.$$

Many optimization problems require to minimize a convex cost function. In other cases, we want to maximize a concave utility function. Maximizing the function f over a given set C is equivalent to minimize the function $-f$ over the same set. If f is concave, $-f$ is convex. The corresponding theorem for the maximization of a concave function can then be immediately derived applying Theorem 1 to $-f$. If we maintain the same definition for the Lagrangian function as in Eq. (2), it holds:

Theorem 2 (Concave optimization with linear constraints)**.** *Consider the optimization problem*

$$\begin{aligned}
\underset{\mathbf{x} \in \mathbb{R}^n}{\text{maximize}} \quad & f(\mathbf{x}) \\
\text{subject to} \quad & \mathbf{a}_i^T \mathbf{x} = b_i, \ \ i = 1, \dots, m \\
& \mathbf{c}_j^T \mathbf{x} \leq d_j, \ \ j = 1, \dots, r \\
& \mathbf{x} \in X,
\end{aligned} \tag{4}$$

where f is concave and continuously differentiable and X is a polyhedral set, i.e. a set specified by the intersection of a finite number of closed half spaces. Then $\mathbf{x}^*$ is a global maximum for this optimization problem if and only if there exist λ_i^*, $i \in \{1, 2, \dots m\}$ and μ_j^*, $j \in \{1, 2, \dots r\}$ such that:

1. *$\mathbf{x}^*$ is feasible,*
2. *$\mu_j^* \leq 0$ and $\mu_j^*(\mathbf{c}_j^T \mathbf{x}^* - d_j) = 0$, $\forall j \in J$,*
3. *$\nabla_{\mathbf{x}} L(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*)^T (\mathbf{x} - \mathbf{x}^*) \leq 0, \ \ \forall \mathbf{x} \in X$.*

2.1 Lagrange multipliers as shadow prices

In specific contexts the Lagrange multipliers often represent quantities with concrete physical meaning. For example, in economic applications they can often be interpreted as prices. In particular, if $f(\mathbf{x})$ is a cost, the Lagrange multiplier associated to a constraint can be viewed as the instantaneous change, per unit of the constraint, in the optimal cost obtained by relaxing the constraint. In other words, it is the marginal utility of relaxing the constraint, or, equivalently, the marginal cost of strengthening the constraint. We are going to show it through a specific simple example, but the result is more general.

Example 1. *Consider the following cost minimization problem*

$$\begin{array}{ll} \underset{\mathbf{x}\in\mathbb{R}^n}{\text{minimize}} & f(\mathbf{x}) \\ \text{subject to} & \mathbf{a}^T\mathbf{x} = b \end{array}$$

where $f(\mathbf{x})$ *is a convex cost function.*[1] *The Lagrangian function is* $f(\mathbf{x}) + \lambda(\mathbf{a}^T\mathbf{x} - b)$ *and Theorem 1 states that if* $\mathbf{x}^*$ *is a global minimum then it exists* λ^* *such that*

$$\nabla f(\mathbf{x}^*) + \lambda^*\mathbf{a} = 0 \text{ and } \mathbf{a}^T\mathbf{x}^* = b,$$

and the optimal cost is $f(\mathbf{x}^*)$*. If the level of the constraint is changed to* $b+\Delta b$*, a new optimum solution* $\mathbf{x}^* + \Delta\mathbf{x}^*$ *is determined with* $\mathbf{a}^T(\mathbf{x}^* + \Delta\mathbf{x}^*) = b + \Delta b$*, and then* $\mathbf{a}^T\Delta\mathbf{x}^* = \Delta b$*. The corresponding cost is*

$$f(\mathbf{x}^* + \Delta\mathbf{x}^*) \approx f(\mathbf{x}^*) + \nabla f(\mathbf{x}^*)^T\Delta\mathbf{x}^* = f(\mathbf{x}^*) - \lambda^*\mathbf{a}^T\Delta\mathbf{x}^* = f(\mathbf{x}^*) - \lambda^*\Delta b.$$

It follows that the sensitivity of the cost to the constraint is:

$$\frac{\Delta\text{cost}}{\Delta b} = \frac{f(\mathbf{x}^* + \Delta\mathbf{x}^*) - f(\mathbf{x}^*)}{\Delta b} = -\lambda^*.$$

Consider for example that $\mathbf{x}$ *represents a vector of quantities of different resources that can be invested in order to reduce the cost* f*, and that the constraint* $\mathbf{a}^T\mathbf{x} = b$ *represents the total amount of resources that can be invested. Now, if less resources are invested (*$\Delta b < 0$*), the total cost increases (*$\Delta\text{cost} > 0$*), and the multiplier* λ^* *is positive. The multiplier represents the cost increase per unit of resource removed and then the implicit value of a resource in the framework of the cost minimization problem. For this reason the multipliers are often referred to as* shadow prices*.*

References

[1] Dimitri P. Bertsekas, Nonlinear programming. *Athena Scientific*, 1999.

[1] Convexity is not required, but we maintain this hypothesis because we only stated some results for Lagrange multipliers in the framework of convex optimization problems.